



Implementation of K-Means Clustering to Group Poverty Levels in Riau Province

Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau

**Nurhafiza Sepriyanti¹, Rahma Sani Nahampun², Muhammad Hafis Zikri³,
Isnani Ambarani⁴, Akhas Rahmadeyan⁵**

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau

⁵Puzzle Reaserch Data Technology, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau

E-Mail: ¹12050321863@students.uin-suska.ac.id, ²12050321674@students.uin-suska.ac.id,
³12050313115@students.uin-suska.ac.id, ⁴12050327049@students.uin-suska.ac.id,
⁵11950314479@students.uin-suska.ac.id

Corresponding Author: Akhas Rahmadeyan

Abstract

The global problem that is being faced and is of concern throughout the world is poverty. Statistically, the poverty rate from year to year sometimes increases and decreases in each region. Based on data from the Pekanbaru Central Statistics Agency, the poverty rate in Riau Province is low at the national level. This is evidenced by the percentage of the poverty rate in Riau in March 2021 of 3.6%. This study aims to classify the poverty level in Riau Province by utilizing one of the existing methods in data mining, namely K-Means Clustering. The data in this study were taken from the Pekanbaru Central Statistics Agency from 2019-2021 which consisted of 12 regencies/cities using the attributes of 3 variables, namely the number of poor people (thousands of people), the number of unemployed (thousands of people) and the average length of schooling (years). The data is then processed using RapidMiner software by classifying the poverty level into 3 clusters, namely: low poverty level (C1), moderate poverty level (C2), and high poverty level (C3). The regencies/cities in cluster 1 are Kuantan Singingi, Indragiri Hulu, Siak, Bengkalis, and Dumai. Then the regency/city in cluster 2 is Pekanbaru. Meanwhile, the regencies/cities included in cluster 3 are Indragiri Hilir, Pelalawan, Kampar, Rokan Hulu, Rokan Hilir, and the Meranti Islands.

Keyword: Clustering, Data Mining, K-Means, Poverty, Riau,

Abstrak

Permasalahan yang dihadapi dan mengkhawatirkan dunia adalah kemiskinan. Secara statistik, tingkat kemiskinan dari tahun ke tahun terkadang mengalami peningkatan dan penurunan di setiap wilayah. Berdasarkan data dari Badan Pusat Statistik Pekanbaru, angka kemiskinan di Provinsi Riau tergolong rendah di tingkat nasional. Hal ini dibuktikan dengan persentase tingkat kemiskinan di Riau pada September 2021 sebesar 7,00%. Penelitian ini bertujuan untuk mengelompokkan tingkat kemiskinan di Provinsi Riau dengan memanfaatkan salah satu metode yang ada pada data mining yaitu K-Means Clustering. Data pada penelitian ini diambil dari Badan Pusat Statistik Pekanbaru dari tahun 2019-2021 yang terdiri dari 12 Kabupaten/Kota dengan menggunakan atribut 3 variabel yaitu jumlah penduduk miskin (ribu jiwa), jumlah pengangguran (ribu jiwa) dan rata-rata lama sekolah (tahun). Data tersebut kemudian diolah menggunakan software RapidMiner dengan mengelompokkan tingkat kemiskinan menjadi 3 cluster, yaitu: tingkat kemiskinan rendah (C1) dengan hasil persentase 2,9%, tingkat kemiskinan sedang (C2) 0,6%, dan tingkat kemiskinan tinggi (C3) 3,5%. Adapun Kabupaten/kota pada cluster 1 yaitu Kuantan Singingi, Indragiri Hulu, Siak, Bengkalis, dan Dumai. Kemudian Kabupaten/kota pada cluster 2 yaitu Pekanbaru. Sedangkan Kabupaten/kota yang termasuk pada cluster 3 yaitu Indragiri Hilir, Pelalawan, Kampar, Rokan Hulu, Rokan Hilir, dan Kepulauan Meranti.

Kata Kunci: Clustering, Data Mining, K-Means, Kemiskinan, Riau

1. PENDAHULUAN

Permasalahan yang dihadapi dan mengkhawatirkan dunia adalah kemiskinan. Saat ini, negara-negara yang kurang beruntung justru menghadapi masalah distribusi pendapatan dan pembangunan yang tidak konsisten, untuk sementara banyak negara non-industri yang mengalami perkembangan finansial yang tinggi namun tidak memberikan keuntungan bagi masyarakat miskin [1]. Indonesia adalah negara non-industri, namun kebutuhan masih menjadi masalah yang harus dipikirkan. Kemiskinan adalah suatu kondisi di mana individu tidak dapat mengatasi masalah-masalah penting seperti makanan, pakaian, pelatihan, dan berbagai kebutuhan dasar lainnya [2]. Kemiskinan juga diartikan suatu keadaan seseorang atau sekelompok orang yang yang tidak mampu memenuhi hak-hak dasarnya untuk mengembangkan dan mempertahankan kehidupan yang baik dan layak [3]. Bagian penting dari teknik pengurangan kebutuhan pendukung adalah aksesibilitas informasi kemiskinan yang tepat dan ditentukan [4].

Provinsi Riau memiliki permasalahan salah satunya yaitu kemiskinan. Selama 3 tahun terakhir (2019-2021), kemiskinan di Provinsi Riau rata-rata mengalami ketidakstabilan. Berdasarkan informasi yang diperoleh Badan Pusat Statistik (BPS) dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) tingkat kemiskinan di Riau pada September 2019 mencapai 6,90%, September 2020 sebesar 7,40% dan September 2021 mencapai 7,00% [5].

Penanganan yang berbeda untuk setiap kabupaten / kota, tergantung pada situasinya, adalah cara yang baik untuk mengatasi kemiskinan. Salah satu caranya adalah dengan mengelompokkan karakteristik masing-masing daerah berdasarkan indikator kemiskinan. Pengelompokan kondisi ini dapat dilihat dalam beberapa aspek. Jumlah penduduk miskin yang tinggal di suatu daerah selama masa pendidikan dan jumlah penduduk yang mengalami pengangguran yang berdasarkan informasi yang diperoleh BPS. Sehingga nantinya dapat diketahui cluster yang sangat berpengaruh pada masyarakat kemiskinan [6].

Adapun penelitian sebelumnya yang pernah dilakukan oleh Yunita, dkk (2020), dalam penelitiannya yang berjudul "Penerapan algoritma K-Means pada klaster data kemiskinan Provinsi Banten menggunakan RapidMiner" menyimpulkan bahwa terdapat tiga klaster. Kabupaten Tangerang, Kota Cilegon, dan Kota Serang termasuk dalam Klaster 0, Kabupaten Pandeglang, Kabupaten Lebak, dan Kabupaten Serang termasuk dalam Klaster 1, serta Kota Tangerang dan Kota Tangerang Selatan termasuk dalam Klaster 2 [7]. Kemudian penelitian Disty, dkk (2019) meneliti tentang "Mengelompokkan Garis Kemiskinan menurut Provinsi Menggunakan Algoritma K-Medoids" serta hasil yang diperoleh yaitu Indonesia memiliki tingkat kemiskinan tinggi yang terdiri dari 23 provinsi dan kemiskinan rendah terdiri dari 11 provinsi [8].

Berdasarkan permasalahan di atas, penulis antusias melakukan penelitian untuk menghasilkan informasi kemiskinan di setiap kabupaten yang ada Provinsi Riau. Untuk menentukan tingkat kemiskinan di suatu wilayah, perlu dilakukan suatu metode untuk mempermudah klasifikasi wilayah yang masih memiliki tingkat kemiskinan yang tinggi. Salah satu metode data mining yang digunakan dalam penelitian ini adalah metode *clustering*. Proses metode ini adalah mengelompokkan nilai yang memiliki kemiripan yang sama pada satu cluster dan menggambarkan pola dari sebuah nilai.

Adapun tujuan dari penelitian ini adalah untuk mengkarakterisasi setiap kabupaten di Provinsi Riau yang memiliki tingkat kebutuhan yang rendah, sedang dan tinggi agar pemerintah dapat mengambil kebijakan dan strategi pembangunan yang tepat sesuai kondisi setiap kabupatennya. Penelitian ini akan dilaksanakan di Provinsi Riau berdasarkan data yang telah diambil dari BPS Provinsi Riau karena masih mengalami permasalahan kemiskinan sehingga membutuhkan campur tangan pemerintah, dunia usaha maupun masyarakat untuk meminimalisir tingkat kemiskinan di Riau.

2. METODOLOGI PENELITIAN

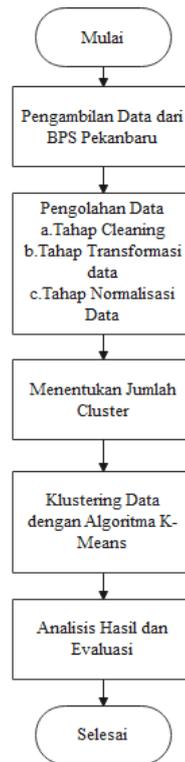
Data yang digunakan diperoleh dari dokumen yang menggambarkan jumlah penduduk miskin di Kabupaten/Kota Provinsi Riau yang dihasilkan oleh publikasi Statistik Indonesia yang diolah BPS. Data yang digunakan adalah dari tahun 2019 hingga tahun 2021. Data yang diperoleh diolah, kemudian di *cluster* menggunakan algoritma K-Means, dan dibagi menjadi 3 *cluster*. Langkah-langkah pada penelitian ini dapat dilihat pada Gambar 1.

2.1 Pengambilan dan Pengolahan Data

Pengambilan data pada penelitian dilakukan untuk mendapatkan informasi penting agar tujuan penelitian dapat tercapai [9]. Data yang kami gunakan didasarkan pada dokumen yang menggambarkan jumlah penduduk miskin di Provinsi Riau di wilayah regional/perkotaan, yang disusun oleh publikasi statistik Indonesia dan diolah oleh Badan Pusat Statistik. Data yang digunakan adalah dari tahun 2019 hingga 2021. Data yang telah didapat akan diolah kemudian setelah itu dilakukan proses *clustering*. Tujuan dari pengolahan data yaitu agar informasi menjadi lebih akurat dan dapat diterima orang lain [10]. Pada tahap ini dilakukan proses penghitungan rata-rata dari setiap data penggunaan untuk mengklasifikasikannya menjadi 3 kriteria penilaian, yaitu:

1. Rata-rata penduduk miskin masyarakat kabupaten/kota di Provinsi Riau
2. Rata-rata jumlah pengangguran masyarakat kabupaten/kota di Provinsi Riau.

3. Rata-rata lama sekolah masyarakat kabupaten/kota di Provinsi Riau.



Gambar 1. Riset Metodologi

2.2 Clustering

Clustering merupakan teknik dalam data mining yang berguna untuk mengelompokkan sekumpulan objek ke dalam beberapa *cluster* dengan karakteristik yang sama, sehingga objek sebuah *cluster* mirip tetapi tidak mirip dengan objek dalam *cluster* yang berbeda. Cluster adalah kumpulan objek yang serupa tetapi berbeda dalam suatu grup dengan benda-benda milik kelompok lain. Ada dua metode *clustering*, yaitu hirarkis *clustering* dan pengelompokan partisi. Data dikelompokkan berdasarkan grafik hierarkis dalam metode pengelompokan hierarkis, dimana dua grup terdekat digabung atau semua data dikelompokkan menjadi cluster. Dalam pengelompokan partisi pengelompokan data tidak memiliki hierarki apapun, setiap cluster memiliki centroid, tujuannya adalah untuk meminimalkan jarak semua data ke centroid [11].

2.3 K-Means

K-Means adalah sebuah metode pengelompokan data menjadi dua atau lebih dari kelompok [12]. Algoritma K-Means merupakan metode analisis kelompok yang membagi subjek penelitian ke dalam kelompok-kelompok, dimana setiap subjek yang akan diamati berada dalam satu kelompok data dengan rata-rata yang berdekatan satu sama lain. Seperti yang kita inginkan, K digunakan untuk konstanta *clustering* total dan Means berarti mean dari dataset, dalam hal ini sebagai *cluster*, jadi Kmeans *Clustering* adalah metode analisis data yang menggunakan sistem partisi untuk pengelompokan data [13].

K-Means merupakan metode pengelompokan data non-hierarkis bagi data menjadi satu atau satu bentuk lebih banyak cluster atau grup untuk data dengan karakteristik yang sama dikelompokkan dalam *cluster* data identik dan karakteristik kelompok yang berbeda adalah kelompok lain [14].

2.4 Min-Max Normalization

Metode Normalisasi Min-Max adalah sebuah metode pengubah data namun tidak menghilangkan isi data tersebut, sehingga pengolahan data akan menjadi lebih mudah. Cara melakukannya dimulai dari normalisasi data dengan menempatkan data pada skala 0 sampai 1, dengan nilai minimum 0, dan nilai maksimum 1. Metode ini memberikan keseimbangan pada data satu dengan yang lainnya [15]. Adapun rumus dari Min Max *Normalization* yaitu:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

3. HASIL DAN PEMBAHASAN

3.1 *Cleaning Data*

Adapun data yang digunakan adalah data kabupaten yang ada di Provinsi Riau dengan atribut berupa rata-rata penduduk miskin dari tahun 2019-2021, jumlah pengangguran dari tahun 2019-2020, dan rata-rata lama sekolah dari tahun 2019-2021 yang diperoleh dari Badan Pusat Statistik Provinsi Riau yang telah dilakukan proses *cleaning*.

Table 1. Hasil *Cleaning Data*

Kabupaten	Rata-Rata Penduduk Miskin			Jumlah Pengangguran			Rata-rata lama sekolah		
	2019	2020	2021	2019	2020	2021	2019	2020	2021
Kuantan Singingi	31,22	29,34	28,9	7419	8376	3291	8,75	8,59	8,58
Indragiri Hulu	26,66	26,66	27,35	10237	10571	7291	8,39	8,38	8,17
Indragiri Hilir	48,29	44,29	44,61	16401	15782	9685	7,24	7,23	7,22
Pelalawan	45,98	45,88	49,3	11235	14223	5915	8,70	8,50	8,49
Siak	24,49	25,38	25,77	9322	12292	9745	9,86	9,66	9,65
Kampar	66,81	65,3	68,74	21954	23911	16797	9,27	9,26	9,25
Rokan Hulu	72,21	73,35	74,73	15054	13958	7537	8,54	8,39	8,38
Bengkalis	35,83	36,96	37,66	24686	24518	17910	9,70	9,69	9,41
Rokan Hilir	49,8	48,85	51,97	14632	13205	9851	8,26	8,25	8,24
Kepulauan Meranti	49,89	47,1	48,5	5615	7475	4060	7,84	7,70	7,51
Pekanbaru	28,6	30,4	32,73	43865	47521	44503	11,92	11,68	11,43
Dumai	10,95	9,88	10,57	9723	12005	9084	10,14	10,07	9,85

3.2 *Transformasi Data*

Selanjutnya dilakukan transformasi pada data penduduk miskin menurut kabupaten/kota Provinsi Riau dengan menentukan nilai rata-rata setiap atribut dari tahun 2019 sampai 2021.

Table 2. Hasil Transformasi Data

Kabupaten	Rata-rata penduduk miskin (2019-2021)	Rata-rata pengangguran (2019-2021)	Rata-rata lama sekolah (2019-2021)
Kuantan Singingi	29.82	6,362.00	8.64
Indragiri Hulu	26.89	9,366.33	8.31
Indragiri Hilir	45.73	13,956.00	7.23
Pelalawan	47.05	10,457.67	8.56
Siak	25.21	10,453.00	9.72
Kampar	66.95	20,887.33	9.26
Rokan Hulu	73.43	12,183.00	8.44
Bengkalis	36.82	22,371.33	9.60
Rokan Hilir	50.21	12,562.67	8.25
Kepulauan Meranti	48.50	5,716.67	7.68
Pekanbaru	30.58	45,296.33	11.68
Dumai	10.47	10,270.67	10.02

3.3 *Normalisasi Data*

Berdasarkan data yang telah ditransformasi sebelumnya, dilakukan normalisasi data menggunakan metode Min-Max *normalization* dengan *range* nilai dalam normalisasi yaitu antara 0-1.

Table 3. Hasil Normalisasi Data

Kabupaten	Rata-rata penduduk miskin 2019-2021	Rata-rata pengangguran 2019-2021	Rata-rata lama sekolah 2019-2021
Kuantan Singingi	0.31	0.02	0.32
Indragiri Hulu	0.26	0.09	0.24
Indragiri Hilir	0.56	0.21	0.00
Pelalawan	0.58	0.12	0.30
Siak	0.23	0.12	0.56
Kampar	0.90	0.38	0.46
Rokan Hulu	1.00	0.16	0.27
Bengkalis	0.42	0.42	0.53
Rokan Hilir	0.63	0.17	0.23

Kabupaten	Rata-rata penduduk miskin 2019-2021	Rata-rata pengangguran 2019-2021	Rata-rata lama sekolah 2019-2021
Kepulauan Meranti	0.60	0.00	0.10
Pekanbaru	0.32	1.00	1.00
Dumai	0.00	0.12	0.63

3.4 Clustering

Tahapan selanjutnya yaitu melakukan *clustering* tingkat kemiskinan di Provinsi Riau berdasarkan jumlah rata-rata penduduk miskin terbanyak, rata-rata jumlah pengangguran dan rata-rata lama sekolah di Provinsi Riau dengan *tools Rapidminer*. Adapun algoritma yang digunakan yaitu K-Means dengan nilai K yang ditetapkan yaitu 3 untuk menghasilkan 3 *cluster*.

Penggunaan algoritma K-Means membentuk 3 *cluster* dan menghasilkan jumlah kecamatan pada setiap *cluster* yaitu *cluster* 1 sebanyak 5 kecamatan, *cluster* 2 sebanyak 1 kecamatan, dan *cluster* 3 sebanyak 6 kecamatan. Adapun nilai centroid yang dihasilkan dapat dilihat pada Tabel 4.

Tabel 4. Centroid Data

Cluster	Jumlah Kabupaten	Rata-rata penduduk miskin (2019-2021)	Rata-rata pengangguran (2019-2021)	Rata-rata lama sekolah (2019-2021)
1	5	0,244	0,153	0,456
2	1	0,319	1	1
3	6	0,712	0,175	0,227

Berdasarkan nilai centroid maka dapat diketahui bahwa *cluster* 1 merupakan kabupaten dengan ciri rata-rata penduduk miskin rendah, rata-rata pengangguran rendah, rata-rata lama pendidikan sedang. Kemudian *cluster* 2 merupakan kabupaten dengan ciri rata-rata penduduk miskin sedang, rata-rata pengangguran tinggi, rata-rata lama pendidikan tinggi. Terakhir *cluster* 3 merupakan kabupaten dengan ciri rata-rata penduduk miskin tinggi, rata-rata pengangguran sedang, rata-rata lama pendidikan rendah.

Tabel 5. Cluster Berdasarkan Kabupaten

Kabupaten	Rata-rata penduduk miskin (2019-2021)	Rata-rata pengangguran (2019-2021)	Rata-rata lama sekolah (2019-2021)	Cluster
Kuantan Singingi	0,307	0,016	0,317	1
Indragiri Hulu	0,261	0,092	0,244	1
Indragiri Hilir	0,56	0,208	0	3
Pelalawan	0,581	0,12	0,3	3
Siak	0,234	0,12	0,561	1
Kampar	0,897	0,383	0,457	3
Rokan Hulu	1	0,163	0,271	3
Bengkalis	0,418	0,421	0,533	1
Rokan Hilir	0,631	0,137	0,229	3
Kepulauan Meranti	0,604	0	0,102	3
Pekanbaru	0,319	1	1	2
Dumai	0	0,115	0,627	1

Penggunaan *software RapidMiner Studio* selama penelitian ini, menghasilkan sebuah *knowledge* sebagai berikut :

1. Kabupaten/kota yang masyarakatnya memiliki Penduduk yang memiliki rata-rata jumlah penduduk miskin sedang, jumlah pengangguran sedang dan rata-rata lama sekolah sedang adalah: Kuantan Singingi, Indragiri Hulu, Siak, Bengkalis, dan Dumai.
2. Kabupaten/kota yang masyarakatnya memiliki Penduduk yang memiliki rata-rata jumlah penduduk miskin banyak, jumlah pengangguran sedikit dan rata-rata lama sekolah sebentar adalah: Pekanbaru.

Kabupaten/kota yang masyarakatnya memiliki Penduduk yang memiliki rata-rata jumlah penduduk miskin sedikit, jumlah pengangguran banyak dan rata-rata lama sekolah terlama adalah: Indragiri Hilir, Pelalawan, Kampar, Rokan Hulu, Rokan Hilir, dan Kepulauan Meranti.

3.5. Hasil Persentase *Clustering*

Tabel 6. Persentase Masing-Masing *Cluster* Berdasarkan Tingkat Kemiskinan Tahun 2021

<i>Cluster</i>	Persentase
1	2,9%
2	0,6%
3	3,5%

Dari tabel 6, dapat diketahui persentase dari masing-masing *cluster* berdasarkan tingkat kemiskinan tahun 2021 *Cluster* 1 memiliki tingkat kemiskinan sebesar 2,9%, *Cluster* 2 memiliki tingkat kemiskinan sebesar 0,6%, dan *Cluster* 3 memiliki tingkat kemiskinan sebesar 3,5%,

4. KESIMPULAN

Kesimpulan dari penelitian ini yaitu hasil untuk 3 *cluster* diperoleh dengan menggunakan *software RapidMiner* dari data yang diperoleh dari Badan Pusat Statistik kota Pekanbaru. *Cluster* satu (1) tergolong *cluster* tingkat kemiskinan rendah karena memiliki jumlah penduduk miskin yang sedikit, tingkat pengangguran sedikit dan rata-rata lama waktu sekolah (pendidikan) sedang. *Cluster* dua (2) tergolong *cluster* tingkat kemiskinan sedang karena memiliki jumlah penduduk miskin sedang, tingkat pengangguran tinggi dan rata-rata lama waktu sekolah (pendidikan) juga tinggi. *Cluster* tiga (3) tergolong *cluster* dengan tingkat kemiskinan tinggi karena memiliki jumlah penduduk miskin yang terbanyak, tingkat pengangguran sedang dan rata-rata lama waktu sekolah (pendidikan) juga dalam tingkat rendah. Dengan adanya pengelompokan tingkat kemiskinan untuk mengidentifikasi karakteristik tingkat kesejahteraan masyarakat setiap daerah yang telah dilakukan pada penelitian ini, pemerintah dapat segera mengambil kebijakan serta strategi yang tepat agar pembangunan di kota Pekanbaru tepat sasaran dan tepat guna.

REFERENSI

- [1] S. R. Dinata, M. Romus, and Yanti, "Faktor Faktor Yang Mempengaruhi Tingkat Kemiskinan Di Provinsi Riau Tahun 2003-2018," *Al-Iqtishad*, vol. 2, no. 16, pp. 116–137, 2020.
- [2] N. Islami and A. Anis, "Pengaruh Upah Minimum Provinsi, Pendidikan Dan Kesehatan Terhadap Kemiskinan Di Indonesia," *J. Kaji. Ekon. dan Pembang.*, vol. 1, no. 3, pp. 939–948, 2019.
- [3] R. Mustafidah, "Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Kemiskinan dengan C-Means dan Fuzzy C-Means Clustering." Institut Teknologi Sepuluh Nopember, 2017.
- [4] G. Anuraga, "Hierarchical Clustering Multiscale Bootstrap Untuk Pengelompokan Kemiskinan Di Jawa Timur," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 3, no. 1, 2015.
- [5] T. P. Purba and T. Himawan, "Pemenuhan Rumah Layak Huni di Provinsi Riau," *J. Penataan Ruang*, vol. 16, no. 2, p. 98, 2021, doi: 10.12962/j2716179x.v16i2.8572.
- [6] A. R. Utami and A. Afriyudi, "PENERAPAN DATA MINING UNTUK CLUSTERING DATA PENDUDUK MISKIN MENGGUNAKAN ALGORITMA HARDC-MEAN DENGAN RAPID MINER," in *Bina Darma Conference on Computer Science (BDCCS)*, 2021, vol. 3, no. 2, pp. 312–320.
- [7] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, pp. 192–198, 2020.
- [8] D. Wahyuli, H. Handrizal, I. Parlina, A. P. Windarto, D. Suhendro, and A. Wanto, "Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids," in *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, 2019, vol. 1, pp. 452–461.
- [9] S. I. Astuti, S. P. Arso, and P. A. Wigati, "Data Dan Pengumpulan Data," *Anal. Standar Pelayanan Minimal Pada Instal. Rawat Jalan di RSUD Kota Semarang*, vol. 3, no. 1996, pp. 103–111, 2015.
- [10] A. Rusmana, "Pengantar Pengolahan Data dan Informasi".
- [11] H. N. Putri and D. R. S. Saputro, "Clustering Data Campuran Numerik dan Kategorik Menggunakan Algoritme Ensemble Quick ROBust Clustering using linKs (QROCK)," in *PRISMA, Prosiding Seminar Nasional Matematika*, 2022, vol. 5, pp. 716–720.
- [12] R. Adha, N. Nurhaliza, U. Sholeha, and M. Mustakim, "Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia," *SITEKIN J. Sains, Teknol. dan Ind.*, vol. 18, no. 2, pp. 206–211, 2021.
- [13] S. Regina, E. Sutinah, and N. Agustina, "Clustering Kualitas Kinerja Karyawan Pada Perusahaan Bahan Kimia Menggunakan Algoritma K-Means," *J. MEDIA Inform. BUDIDARMA*, vol. 5, no. 2, pp. 573–582, 2021.

- [14] E. Muningsih, I. Maryani, and V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” *EVOLUSI J. Sains dan Manaj.*, vol. 9, no. 1, 2021.
- [15] T. T. Hanifa, S. Al-Faraby, F. Informatika, and U. Telkom, “Analisis Churn Prediction pada Data Pelanggan PT. Telekomunikasi dengan Logistic Regression dan Underbagging,” *vol*, vol. 4, pp. 3210–3225, 2017.